

WHAT IS CLAIMED IS:

1. A method for substantially real-time analyzing of a stream of data comprising:
  - receiving the stream of data;
  - determining a data distribution representative of the stream of data,including creating data bins having exponentially increasing sizes; and
  - allocating statistical representation of the data in the data bins; and
  - using the data distribution to analyze the stream of data.
2. The method of claim 1, wherein creating data bins having exponentially increasing sizes includes indexing the bins using a set of keys determined from a function of the logarithm of the incoming data, determining a set of exponentially increasing intervals to define the data bin sizes.
3. The method of claim 2, wherein determining the set of keys includes defining a resolution factor as a number of data bins desired per power of the chosen logarithm base; and using the resolution factor to determine the set of exponentially increasing intervals.
4. The method of claim 1, wherein receiving the stream of data includes querying a data source and collecting the stream of data from the data source in response to the query.
5. The method of claim 1, comprising defining the data stream as a continuous stream of data having a high data rate.
6. The method of claim 1, comprising defining the stream of data as having only positive values.

7. The method of claim 1, comprising defining the stream of data as having an unknown lowest value and an unknown upper value.
8. The method of claim 1, comprising defining a bin order; and storing the bin order in memory.
9. The method of claim 8, comprising the bin order as an array structure; and storing the data bins in the array structure in memory.
10. The method of claim 9, wherein recording statistical data representative of the incoming data value in the data bins includes receiving a data value; computing a bin key associated with the data value; define an array index having an array of index values wherein each array index value is associated with a data bin; determine the data bin associated with the data value using the array index and bin key.
11. The method of claim 10, further comprising updating the value stored in the data bin.
12. The method of claim 10, wherein if a data bin cannot be determined, extending the array structure to accommodate the data value.
13. The method of claim 9, further comprising indexing the bins using a set of keys.
14. The method of 9, further comprising defining the array structure as a tree array structure.
15. The method of claim 14, wherein allocating a data value in the tree array structure includes determining a data bin for the data value, and if a data bin does not exist, creating a data bin.

16. A system for analyzing a stream of data comprising:  
a dynamic distribution collector configured for receiving the stream of data, and determining a data distribution representative of the stream of data, including configured to create data bins having exponentially increasing sizes, and recording a statistical representation of the data in the data bins.
17. The system of claim 16, wherein the dynamic distribution data collector is configured for indexing the bins using a set of keys determined from a function of the logarithm of the incoming data, and is configured to determine a set of exponentially increasing intervals to determine the data bins sizes.
18. The system of claim 16, wherein the data is usage data.
19. The system of claim 16 wherein the dynamic distribution data collector is configured to order the bins in an array structure.
20. A computer-readable medium having computer executable instructions for performing a method for substantially real-time analyzing of a stream of data comprising:  
receiving the stream of data;  
determining a data distribution representative of the stream of data,  
including creating data bins having exponentially increasing sizes; and  
allocating statistical representation of the data in the data bins; and  
using the data distribution to analyze the stream of data.